焦虑个体趋避冲突失调的认知神经机制

夏熠 张婕 张火垠 雷怡 窦皓然

四川师范大学脑与心理科学研究院，成都 610066

摘 要 在日常生活中，如何有效应对趋避冲突至关重要，而焦虑个体存在趋避冲突失调的行为表现。这种失调表现为以放弃积极结果为代价，以此回避与实际威胁无关或威胁程度较低的刺激。以往的动机理论将个体应对趋避冲突分为信息输入和行为输出过程，难以全面解释趋避冲突失调的具体机制。因此，本文尝试提出一个全新的“冲突感知、冲突处理和反馈学习”三阶段模型，强调焦虑个体趋避冲突失调可能表现为威胁感知的增强、预期价值与动机比较的失衡、反馈学习的异常。未来研究可以进一步验证该模型中三阶段的相对独立性，通过分层和模块化的方法对模型进行参数化，从发展的角度来探讨焦虑个体趋避冲突失调背后的认知神经机制。

关键词 焦虑，趋避冲突，预期价值，动机，认知神经机制

趋近和回避之间微妙平衡的紊乱是包括焦虑在内的许多精神障碍的重要表现，即趋避冲突中异常的行为决策(Aupperle & Paulus, 2010; Vogel & Schwabe, 2019; Letkiewicz et al., 2023)。趋避冲突的概念最早由 Lewin(1935)提出，指当一个目标同时导致积极和消极的结果时，个体既想要追求积极结果又想避免消极结果，从而产生进退两难的心理冲突。比如，社交焦虑患者可能希望多交一些朋友，但是又想回避社交场合所带来的不适感。在此情境中，个体通过整合奖励与威胁的相关信息以及结果发生的可能性，从而做出趋近或回避的行为反应(Quartz, 2009; Kirlic et al., 2017)。回避模型是焦虑障碍的经典模型，但是真实社会环境比较复杂，不仅包含对威胁刺激的回避，也包含对积极刺激的趋近。所以，趋避冲突的实验模型逐渐取代传统的回避范式，成为评估和预测焦虑的综合模型(Korn et al., 2017; Bach, 2022; Ball et al.,2022)。在没有特别说明的情况下，本文中的焦虑个体泛指临床、亚临床以及焦虑倾向的个体。

收稿日期：2024-02-02

* 国家自然科学基金面上项目 (32271142)、教育部哲学社会科学研究重大课题攻关项目 (21JZD063)资助。

通讯作者：窦皓然，E-mail：haorandou@sicnu.edu.cn

一直以来，对于趋避冲突的理解主要从动机维度出发(Gray, 1987; Elliot, 2006; Monni et al., 2020)：个体在趋近回避冲突中主要依赖趋利避害的内驱力进行决策，并且将个体在整个趋避冲突中的反应简化为信息输入和行为输出两个阶段。然而，焦虑个体在奖赏刺激很强且威胁刺激很弱的情况下，依旧选择回避行为 (Arnaudova et al., 2017; Pittig et al., 2021) ，并且只需要更少的信息就能快速地做出反应(Dillon et al., 2022; Liu et al., 2022; Han et al., 2023)。这表明他们的行为可能并不仅是为了避免痛苦或寻求快乐，还可能源于对潜在威胁的过度反应或习惯性回避(Ball et al.,2022; Watson et al., 2022)。当然，焦虑个体趋近回避冲突失调也可能是由于预期价值计算的受损。因为从进化角度来说，个体的决策往往基于价值的最大化，那么这种放弃更高价值的非理性决策行为可能是个体对不同行为对应结果的概率和成本的复杂计算出现了问题(Rangel et al., 2008; Walters & Redish, 2018)。然而现有理论框架未充分考虑预期价值与动机的相互作用，难以全面解释焦虑个体趋避冲突失调的具体机制。因为过度回避可能代表的是回避相关的动机程度，也可能反映的是预期价值计算的受损。同时，趋避冲突的解决涉及多个脑区的协同作用，传统的理论框架将趋近和回避动机作为基本独立的神经机制，局限于独立脑区激活的结果，而焦虑通常与多个脑区的过度激活或功能损伤相关，这些区域并非独立工作，而是通过广泛的网络连接来共同调节行为决策。基于现有理论在认知和神经机制中的不足，本文提出个体应对趋近回避冲突的三阶段模型，我们的核心科学问题是："焦虑个体在冲突感知、冲突处理和反馈学习三个阶段中的认知神经变化，如何共同作用导致趋避冲突中的异常行为？"对此，本文首先总结了基于强化敏感性、巴甫洛夫条件化以及强化学习等与趋避冲突和焦虑相关的理论及其局限性，在以往理论基础上提出"冲突感知、冲突处理和反馈学习"的趋避冲突三阶段模型；其次，总结该模型所依赖的认知神经基础；最后，探讨焦虑在威胁感知增强、预期价值和动机比较失衡以及反馈学习上的异常表现，解释焦虑对不同阶段的影响，为理解焦虑个体的趋避冲突失调提供新的视角，并对未来的研究方向提供建议。

1. 趋避冲突系统的理论基础

1.1 强化敏感性理论

Gray(1987)首次提出强化敏感性理论(Reinforcement Sensitivity Theory, RST) 解释焦虑个体的异常回避和趋近行为。该理论包含三种动机系统：行为激活系统(Behavioral Activation System, BAS)与奖励、积极情感及主动性相关，促进个体继续当前的行为；行为抑制系统(Behavioral Inhibition System, BIS)与惩罚、负面情感和被动性相关，抑制当前行为；风险因

素系统(Fight-Flight-Freeze system, FFFS)对危险产生自动应激反应，包括战斗、逃跑或僵化。在面对趋避冲突时，BIS 作为冲突检测和处理系统，抑制正在进行的行为并唤醒注意资源(Gray & McNaughton, 2000)，焦虑状态被认为是 BIS 过度激活所引起(Corr & Cooper, 2016)。在此基础上，研究者提出了引导动机行为的两阶段模型(Corr & McNaughton, 2012; Corr & Cooper, 2016)：在评估输入阶段，个体识别到强化物（刺激）同时具有吸引和排斥的属性，引起矛盾心理；在动机输出阶段，相应的行为系统（如 BAS、FFFS 或 BIS）被激活，产生趋近和回避的动机并做出最终决策。该模型解释了冲突发生时的输入和输出阶段，并强调了动机在理解和处理趋避冲突中的核心作用。然而强化敏感性理论的初衷在于解释特定刺激下的行为选择，但对于焦虑个体，相同的回避行为可能在多种不同刺激下重复发生，显示出普遍的回避偏向性。

对此，Corr 和 McNaughton(2012) 扩展了强化敏感性理论的特质维度用以解释焦虑的稳定性，并使用不同的敏感性来进行标记。其中，BAS 的敏感性被定义为奖赏敏感性(Corr, 2004)，BIS 的敏感性被定义为威胁敏感性(李小新 等, 2014)。焦虑个体容易受威胁敏感性的驱动从而引起过度的回避行为(Corr & Krupić, 2017)，而奖赏敏感性可能随着焦虑的发展逐渐降低，从而减少趋近行为 (Richey et al., 2019; Sequeira et al., 2022)。在强化敏感性理论的框架下，焦虑个体趋避冲突的失调可以被解释为威胁敏感性和奖励敏感性的失衡(Bishop & Gagne, 2018)。虽然强化敏感性理论通过不同程度的敏感性差异影响回避或趋近动机来解释焦虑个体趋避失调的原因，但是该理论过度简化了个体在面对复杂决策时的内部认知过程，也忽略了动机和认知之间的相互作用。有效地认知处理能够帮助个体平衡趋近和回避动机，但是这种认知处理本身可能在焦虑个体中出现异常，因此对于趋避冲突的解释还需要从认知过程的维度进行扩展。

1.2 巴甫洛夫条件反射和回避学习理论

巴甫洛夫恐惧条件反射是研究焦虑的经典模型。行为主义认为回避是通过反复的刺激-反应联结所产生，因此回避行为是对焦虑和恐惧的直接反应(Watson & Rayner, 1920; Mowrer, 1940; Miller, 1948) 。根据这一观点，回避行为将直接减少面对的恐惧，从而被不断强化。然而，许多回避行为并未伴随显著的恐惧减弱，这表明减少恐惧可能不是唯一的驱动力(Bolles, 1970; Rachman & Hodgson, 1974)。于是研究者逐渐将目光转移到认知层面，关注威胁预期在回避行为的习得和维持中的作用（即虽然恐惧减弱，但对威胁的预期并没有减弱），将回避从被动的行为反应转向具有认知驱动的主动行为(Seligman & Johnston, 1973)。Lovibond(2006)提出了预期模型，认为回避不仅是条件反射的结果，还涉及个体对结果的预

测和认知评估，该模型强调威胁预期是回避行为的核心驱动力。另一种对持续性回避行为的解释是习惯化，即回避行为的高度自动化(Krypotos et al., 2015; Hofmann & Aleena, 2018)。因此，目前普遍认为回避行为由两个相互独立的学习过程共同支配(de Wit & Dickinson, 2009; LeDoux et al., 2017; Watson et al., 2022)。首先是威胁预期过程：这是一种目标导向的学习，即个体关注威胁线索和结果之间的联系，其回避反应是出于对威胁结果出现可能性的估计(de Wit & Dickinson, 2009)。焦虑个体在面对威胁线索时可能会高估威胁结果出现的可能性，从而产生较高的恐惧水平和过度的回避反应；其次是习惯化过程：通过大量重复的行为练习，个体的回避反应逐渐自动化，对威胁线索的回避反应成为习惯，最终使得回避行为脱离了对结果的依赖(Wood & Neal, 2007)。虽然经典条件反射和回避学习理论在一定程度上补充了强化敏感性理论所缺乏的认知加工过程，并解释了在得到结果反馈后，存在两种不同的学习系统应对相同或类似的情况，但是该理论只探讨了威胁情况下个体如何从刺激-行为-结果的联结中进行学习，并不能全面地解释焦虑个体趋避冲突失调的机制。

1.3 强化学习理论

强化学习理论(Reinforcement Learning Theory) 为奖励驱动(或威胁驱动)行为的建模提供了一个框架，个体在此框架下通过权衡决策结果的预期值，旨在最大化奖励并最小化损失。强化学习的概念源自操作性条件反射，行为主义者认为通过强化物可以有效塑造行为，而无需推测个体的内部心理过程(Skinner, 1938) 。然而，Tolman(1948)发现动物可以通过对环境形成认知地图，在没有强化物的情况下也能进行学习并调整行为。这一发现促使研究者从单纯依赖强化物的行为主义转向了包含预测、计划和决策的认知心理学框架(Silvetti & Verguts, 2012)。Rescorla-Wagner(1972)模型提出预期结果和实际结果之间的差异（预测误差）是驱动学习的关键。随后，Sutton(1988)将预测误差引入时序差分学习(Temporal-Difference Learning)，即个体在每一步都更新对未来状态和结果的预期值，进而描述个体如何通过不断更新对环境的预期来进行适应性的行为调整。发展至今，强化学习通过计算预期价值来指导行为决策，其中奖励值代表趋近行为的预期收益，而惩罚值则表示某行为可能导致的负面后果。预测误差被认为是调整和优化行为决策的核心因素：如果行为的结果高于预期，个体会增强对该行为的倾向；如果实际结果不如预期，个体将减少相同行为的发生(Sutton & Barto, 2018; Sharp et al., 2022; Letkiewicz et al., 2023; Enkhtaivan et al., 2023)。利用计算建模，该理论揭示了个体如何动态适应环境和调整预期的内部认知过程。

在此基础上，焦虑个体异常的行为决策被认为是对学习、感知和决策方面计算的异常变化所导致(LeDoux & Pine, 2016; Bach & Dayan, 2017; Yamamori & Robinson, 2023)。Pike(2022)

对强化学习参数的元分析发现相比于对照组，焦虑障碍患者对惩罚学习率的升高和奖励学习率的降低，并将其解释为焦虑障碍对负面信号的高敏感性。该理论框架通过量化奖励和威胁的期望值，可以更深入地理解个体进行权衡和决策的认知过程。然而，强化学习理论容易将预期和内部动机混淆。因为预测错误的定义是结果与预期之间差异的符号（正或负），这种定义将预测错误的大小和结果的动机性质混合，例如，超出预期的奖励（正的预测错误）通常被视为有吸引力，而超出预期的威胁（也是正的预测错误）则被视为有厌恶性。相反，比预期少获得的奖励被视为厌恶性，而比预期少遇到的威胁则被视为有吸引力(Kalisch et al., 2019; Moughrabi et al., 2022)。传统的强化学习模型通常不擅长处理多维度结果，这些后果难以简单地分类为正面或负面，因此限制了其在解释复杂情境下个体如何权衡趋避冲突的适用性。

1.4 趋避冲突的三阶段模型

综上所述，我们认为前人的动机理论将趋避冲突的处理简单划分为一个静态的评价输入和动机输出的两阶段模型，难以解释内部的认知神经机制和个体根据结果进行适应性的调整过程。而描述了认知过程和反馈学习的理论又往往局限于奖赏或威胁的单一维度，没有纳入趋避冲突的情况。因此，我们对上述理论基础进行了整合和扩展，强调冲突处理过程中价值和动机以及两者的交互作用，并且在得到反馈后获得强化或者更新，从而形成一个动态的认知回路（如图 1）。具体而言，当冲突刺激输入时，引起个体对刺激的注意和解释，这受到威胁和奖赏敏感性的影响。随后是冲突处理阶段，个体会对回避和趋近行为分别产生预期价值和动机，并将两者进行比较后做出决策。最后是反馈学习阶段，如果结果反馈与预期不相符，个体以目标导向的学习来调整对信息的判断；如果结果反馈与预期相符，个体强化现有模式，对相同或类似的信息形成习惯化从而维持一贯的行为。下文将对该三阶段模型进行详细介绍，阐明该模型的合理性与独特性。

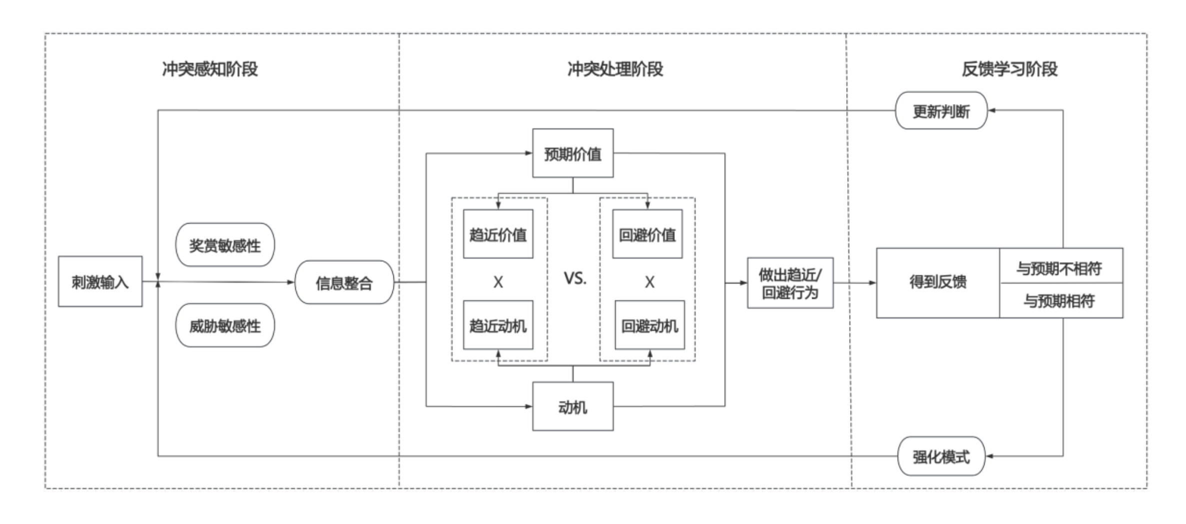


图 1. 趋避冲突的三阶段模型

2. 趋避冲突三阶段模型的认知神经基础

2.1 趋避冲突的认知机制

2.1.1 冲突感知

当个体面临趋避冲突刺激时，会激活强化敏感性系统进行初步的感知和分类(Corr & McNaughton, 2012; Corr & Cooper, 2016; Monni et al., 2020)。冲突感知阶段涉及的是对冲突刺激的快速评估和响应，即强化敏感性系统的奖赏敏感性和威胁敏感性，它们各自影响个体对刺激的注意力分配和理解。例如，高威胁敏感性的个体可能对威胁刺激产生较强的生理唤醒和持续的注意，这通常表现为对威胁线索过度敏感，从而在遇到潜在威胁时迅速产生回避行为(De Pascalis et al., 2019)；相对而言，高奖赏敏感性的个体则可能对奖励刺激的识别和反应更加敏感，使得他们在预测到潜在的积极结果时更倾向于趋近行为(Amodio & Jones, 2011; Kaye et al., 2018)。敏感性差异还影响了个体对外部事件的理解，Hundt(2013)要求被试连续一周每天 8 次汇报自己日常生活中的行为、感受和思考，发现高威胁敏感性的个体在日常生活中经历的负面情绪较多，而在积极情境中感受到的正面情绪增加较少，反映出消极的解释偏向。他们也更倾向于将环境线索视为潜在威胁，从而激发更多的消极情绪，而高奖赏敏感性个体则表现出相反的倾向(Warr et al., 2021)。虽然健康个体通常能够较好地平衡这两种敏感性，但敏感性的失衡可能构成精神障碍的基础(Bishop & Gagne, 2018； Katz et al., 2020)。基于威胁敏感性和奖赏敏感性的差异，个体在感知趋避冲突阶段可能已经产生对回避或趋近行为的偏好，并进一步影响下一个阶段。

2.1.2 冲突处理

个体对信息进行整合后，开始处理趋避冲突并做出决策，这个过程需要动机和预期价值的共同作用(Roesch & Olson, 2004; Verharen et al., 2020)。在趋避冲突的情况下，追求奖励的趋近动机与避免威胁的回避动机相互对立，产生动机竞争 (McNaughton et al., 2016; McNally, 2021)。其中，趋近动机通常由行为激活系统(BAS)控制，驱使个体追求可能带来积极结果的行为；而回避动机主要涉及避免惩罚和威胁，由风险因素系统(FFFS)和行为抑制系统(BIS)控制，促使个体避开可能产生消极结果的选择。趋近和回避动机的强度因人而异，这被认为是对某种行为的主观偏好(Gray & McNaughton, 2000; Corr, 2004; Corr & Cooper, 2016)。除此之外，个体还会基于对行为与结果之间关系的推测，从客观角度计算趋近和回避的预期价值(Lee et al., 2012; Biderman et al., 2020)。相应地，研究发现个体在面临可能带来奖励或威胁的选择时，与结果相关的记忆表征变得活跃，并且表征的强度预测了随后的行为选择， 佐证

了在做出决定前存在一个对潜在结果的价值进行模拟的过程(Castegnetti et al., 2020; Cisler et al., 2023)。尽管面对威胁与损失奖赏的预期编码在认知过程中具有相似性(Tom et al., 2007; Kalisch et al., 2019)，但是个体对损失的厌恶通常超过同等大小的收益(Kahneman & Tversky, 1979; Tymula et al., 2023)。这表明获得和损失并不可以等值替换，获得奖励和面对威胁共同决定趋近的价值，而失去奖励和避免威胁共同决定回避的价值。进一步的研究发现，虽然奖励可以调节趋避冲突中的行为反应，却不能调节对威胁的预期(Schlund et al., 2016; Pittig et al., 2018; Pittig & Dehler, 2019)。也就是说趋近和回避的预期价值可能是先独立计算出来，随后再相互比较从而做出决策。

在实际决策过程中，动机和预期价值会相互影响，动机引导个体的趋近或回避倾向，而预期价值反过来又调节动机的强度和方向。具体来说，Moughrabi 等(2022)让被试在三个图案中进行选择，每个图案所获得奖励（积分）和遭受威胁（电击）的概率都不同，每个试次都包括了三个阶段：选择、预期和结果，并分为一致和冲突两种条件，在一致条件下，奖励高的图片触发电击概率的可能性低；而在冲突条件下，奖励高的图片发生电击的概率也高。行为结果发现，被试在较高的威胁情况下仍然倾向于趋近行为，表现出明显的偏好。模型比较显示，将威胁结果和奖赏结果的价值整合到一个总体期望中的模型效果更好，这些价值由个体对威胁和奖励的偏好加权。换而言之，个体先对回避和趋近行为的预期价值与动机进行乘积后，再进行比较和整合。并且冲突的核心不仅是刺激的动机属性，而更多地涉及对未来奖励或威胁的预期(Doll et al., 2015; Wise et al., 2021)。经过对预期价值和动机的综合评估后，个体做出回避或趋近行为并得到结果反馈。

2.1.3 反馈学习

结果反馈会影响未来的决策(Miletić et al., 2021)，如果结果与预期相符合，个体会强化现有的行为模式，奖励结果有助于强化刺激-行为的关联，从而增加趋近行为的可能性(Ranaldi., 2014)。威胁结果则被编码以避免未来类似的威胁，从而加强回避行为(Feigley & Spear, 1970)，最终形成习惯化；如果结果与预期不相符，个体计算预测误差，实际与预期结果间的差异触发目标导向的学习，基于预测误差的大小和方向，重新调整回避或趋近的策略(Diederen & Schultz, 2015)。根据学习的双系统理论，习惯化系统会与目标导向系统进行竞争(Keramati et al., 2011; Wood & Rünger, 2016)。Glück 等人(2021)结合趋避冲突来评估习惯性回避对目标导向学习的影响，首先是习惯性回避阶段，让被试重复对特定刺激做出回避以避免电击，随后移除电击来降低威胁刺激的价值，最后考察被试是否仍然表现出习惯性回避，即便此时选择趋近行为可能带来奖励。结果显示，即使特定刺激后面不再跟随电击而是奖励，

习惯性回避仍然影响着被试的决策，特别是在目标导向的行为与之前习惯化回避的不一致时，准确率明显降低。总之，习惯化行为减少了个体对外部反馈和预测误差的敏感性，节省了认知资源的使用。然而当与预期不一致时，即使明确提示旧有的行为策略不再最优，习惯化行为会干扰目标导向的学习，造成认知和行为上的不一致。同时，无论是哪种反馈学习占据主导，都会在整个认知模型中循环，继而影响后续相同或类似情境的决策。

2.2 趋避冲突的神经机制

2.2.1 冲突感知：情绪动机相关的皮质下结构激活

随着刺激的输入，前扣带回唤醒对冲突刺激的注意，并快速启动情绪驱动的加工过程，这涉及到杏仁核和腹侧纹状体等与情绪动机相关的皮质下结构(Loijen et al., 2020)。对于冲突感知阶段，杏仁核(Amygdala)在感知威胁时迅速响应，促进个体对潜在威胁保持警觉，从而增强对威胁刺激的注意和解释偏向，使高威胁敏感性个体的加工速度更快，注意力更集中(Choi & Kim, 2010; Diehl et al., 2019; Miller et al., 2019)；相反，腹侧纹状体(ventral striatum)的激活增强了个体对奖赏刺激的注意和动机，从而促进趋近动机并抑制回避动机，高奖赏敏感性个体表现出更大的奖赏关注和积极的解释偏向(Nguyen et al., 2019; LeBlanc et al., 2020)。如图 2a，在面对趋避冲突刺激时，首先被激活的是情绪和动机驱动的自下而上的加工系统，主要涉及对威胁敏感的杏仁核以及对奖赏敏感的纹状体，这些结构为刺激提供情绪效价并传递信号。

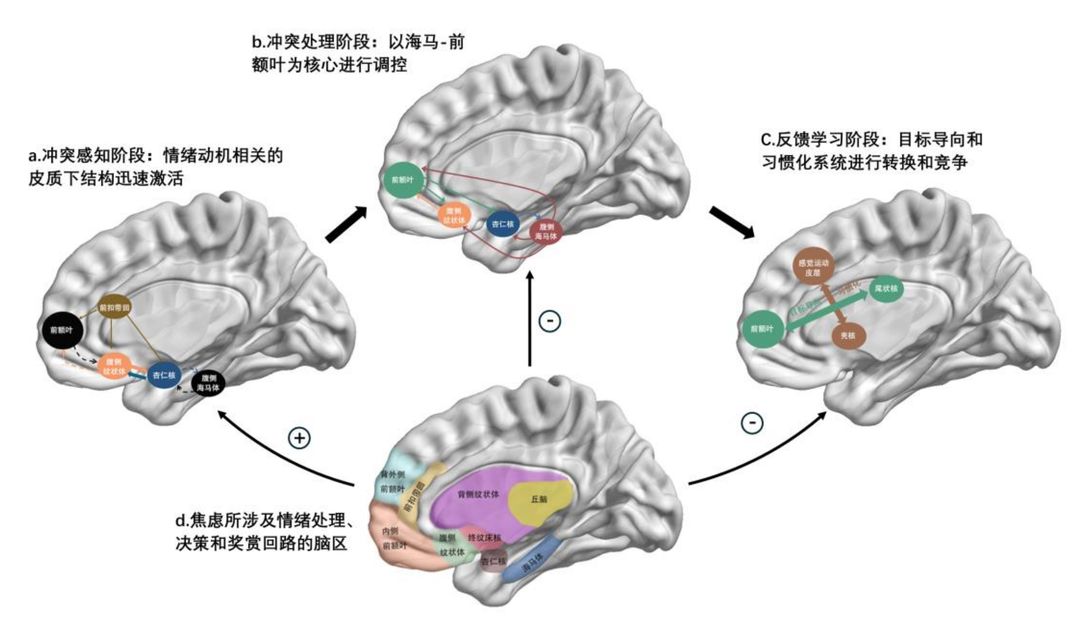


图 2. 焦虑影响趋避冲突失调的神经机制。(a) 冲突感知阶段主要涉及前扣带回的注意分配以及纹状体和杏仁核的快速激活；(b)冲突处理阶段涉及海马对趋近回避的比较以及前额叶自上而下的调控和编码预期；(c)反馈学习阶段涉及皮质和基底神经节的不同回路之间

的转换和竞争；(d)焦虑涉及的脑区，主要是增强了杏仁核的活动以及损坏了前额叶和海马的功能，导致对威胁注意的增加和预期失衡，以及学习系统的转化不灵活。

2.2.2 冲突处理：以海马-前额叶为核心的调控通路

在自下而上的快速加工完成后，冲突处理阶段逐渐由自上而下的加工主导 (Kelley et al., 2017; Lacey & Gable, 2020; McNally, 2021)。此时，个体需要编码预期并比较趋避冲突，从而做出决策。这个过程主要依赖于海马和前额叶皮层及其之间的通路。

腹侧海马(Ventral Hippocampus, vHPC)被认为是处理冲突刺激的关键区域，并与动机行为密切相关(Ito & Lee, 2016; Bryant & Barker,2020; Fernández-Teruel & McNaughton, 2023)。在高度趋避冲突的情况下，vHPC 激活并发挥看似相互排斥的作用。一方面 vHPC 抑制对动机冲突刺激的趋近反应(Rusconi et al., 2022)。例如，Bach 等(2014)要求被试在被劫匪抓住的威胁下选择收集金币(趋近)或为避免被抓住导致失去所有金币(回避)，并设置了 3 种威胁等级，结果发现腹侧海马的同源区在受到威胁时激活以阻止趋近行为。进一步的研究表明，vHPC 的损伤会增加冲突下的趋近行为(Bach et al., 2019)，回避行为也更能解释海马的血氧水平依赖信号(Abivardi et al., 2020)。另一方面，vHPC 亚区的失活也导致对冲突刺激的整体回避，表明这些区域对趋近行为施加独立的控制，并促进趋近行为(Schumacher et al., 2018; Yeates et al., 2022)。这些证据说明 vHPC 在趋避冲突中表现出分离作用，可能是将回避和趋近进行比较的关键区域。前额叶皮层(prefrontal cortex，PFC)是自上而下的认知控制的主要区域，其中内侧前额叶(medial prefrontal cortex, mPFC)是处理趋避冲突的关键节点(Duval et al., 2015)。不同的 mPFC 亚区和神经元群通过激活不同的下游回路，编码奖励和威胁刺激(Ye et al., 2016; Rozeske et al., 2018; Pastor et al., 2020)，并参与奖励和厌恶体验(Del Arco, 2020)。Gazit(2020)通过记录被试在趋避冲突任务中 mPFC 以及杏仁核和海马的神经元活动，发现 mPFC 神经元更倾向于对威胁做出反应，其选择性编码动机结果的消极性。这些研究说明 mPFC 不仅能调控对奖励和威胁刺激的响应，还是预期编码的关键区域。

此外，趋避冲突的处理还涉及其他神经回路，其中一些倾向于追求奖励，而另一些则更多地避免威胁。vHPC 对趋避冲突进行信息整合并做出决策，既要接受来自包括杏仁核、纹状体的信息输入，还需要将信息输出到前额叶皮层等区域，从而形成网络连接(如图 2b.)。其中 vHPC-杏仁核(amygdala, AMY)和 vHPC-伏隔核 (nucleus accumbens, NAcc) 作为主要的传入回路，vHPC-AMY 对整合恐惧相关刺激尤为重要(Yeates et al., 2020；Bryant &Barker., 2020)。vHPC-NAcc 回路与对线索的决策以及维持动机行为有关(Barker et al., 2018)，该回路的化学抑制会导致决策时间增加(Patterson & Arruda, 2020)。mPFC-vHPC 则作为一个门控系统，PFC

可以使用 vHPC 的信息来指导认知控制，从而编码奖赏信号以促进下游脑区制定趋近或回避的策略(Padilla-Coreano et al., 2016; Moscarello & Maren, 2018)。mPFC 对杏仁核自上而下的调节被厌恶刺激激活，从而产生恐惧反应并促进回避行为(Kirry et al., 2020; Fernández-Teruel & Tobeña, 2020)；同时，mPFC 对 NAcc 的投射则被奖励刺激激活，并促进趋近行为(Ma et al., 2020)，该通路涉及对奖励相关刺激的激活(Otis et al., 2017)，以及在面对厌恶线索时抑制寻求奖励的作用(Piantadosi et al., 2020)。这些结构和通路共同参与精细的神经调节，从而决定个体对奖励和威胁刺激的反应策略。

2.2.3 反馈学习：皮质-基底神经节中的竞争通路

通过自上而下对趋避冲突的比较以及价值编码后，个体得以做出决策并开始反馈学习。反馈学习依赖的是连接皮质和基底神经节(basal ganglia, BG)回路(cortical-BG)，纹状体在其中起重要作用，可以自主计算用于学习的预测信号误差，然后再将信息传入回路从而调整行为策略(Engelhard et al., 2019; Baladron & Hamker, 2020)。在这一过程中，存在两种学习系统：目标导向系统和习惯化系统，两个系统在皮质-基底神经节回路中相互竞争。

目标导向系统依赖于前额叶(PFC)和背内侧纹状体(dorsomedial striatum, DMS 相当于人类的尾状核)之间的回路，其根据预期与实际结果之间的差异计算预测误差，使得个体能够灵活地调整行为；习惯化系统则涉及背外侧纹状体(dorsolateral striatum, DLS 相当于壳核)在内的感觉运动皮层-基底神经节回路，通过反复地强化现有模式，可以减少个体对外部反馈和预测误差的敏感性，从而节省认知资源的使用 (Kim & Hikosaka, 2015; Wood & Rünger, 2016; Baladron & Hamker, 2020)。Barnett 等(2023)通过模拟皮质-基底神经节回路，让模型执行追选任务，并测试奖励贬值和反转条件下模型的适应性。结果发现 DLS 主导的学习显示出对新规则的抵抗，表现出习惯化行为；而 DMS 主导的学习则能适应新规则，反映出目标导向的学习。在破坏 PFC 的功能后，目标导向的学习能力降低，习惯更难以被改变。总之，虽然目标导向和习惯化这两种反馈学习系统都依赖于皮质-基底神经节回路，但它们彼此竞争(如图 2c)。长期的习惯化行为可能导致大脑的决策制定路径变得相对固定，因而更新行为策略变得更加困难，即使个体意识到新的策略可能带来更大的奖励，改变根深蒂固的行为模式也需要更多的努力。

3 焦虑个体的趋避冲突失调

趋避冲突的三阶段模型认为解决趋避冲突包括冲突感知、冲突处理和反馈学习（图 1），而焦虑的影响可能不限于某一个阶段，且焦虑涉及的脑区与解决趋避冲突所依赖的神经机制

相重叠(Aupperle & Paulus, 2010; Fernández-Teruel & Tobeña, 2020，图 2d)。接下来，我们将详细探讨焦虑对这三个关键阶段的具体影响，揭示其趋避冲突失调的整体机制。

3.1 威胁感知的增强

研究者开发了行为抑制量表(Behavioral Inhibition Scale, Carver & White, 1994) 作为敏感性测量工具，发现行为抑制得分与焦虑得分呈正相关且可以预测焦虑水平(Izadpanah et al., 2016)。高威胁敏感性的个体会更倾向于关注与威胁相关的线索并产生更消极的反应，这种倾向可能导致长时间的焦虑和恐惧。自我报告和点探测的研究发现，焦虑个体对威胁的注意和解释加工都会受到影响(Mathews & MacLeod, 2002; Bar-Haim et al., 2007; Kreuze et al., 2020)。Yamamori 等(2023) 通过计算建模对威胁敏感性和奖赏敏感性进行参数化，研究结果发现较高的威胁敏感性能解释处于焦虑状态的个体在趋避冲突中为何倾向于采取回避行为。在趋避冲突的早期输入阶段，焦虑个体的高威胁敏感性导致他们对威胁刺激的注意和解释偏差，这种威胁感知能力的增强可能反映出了焦虑个体对潜在的威胁更加敏感，从而倾向于选择回避作为一种适应性的反应。

以威胁、奖赏的神经处理为目标的研究确定了焦虑所带来的影响，主要表现为焦虑使得杏仁核对威胁的反应性增加，也使纹状体对奖赏的反应性降低(Stelly et al., 2020; Chavanne & Robinson, 2021; McDermott et al., 2022; Auerbach et al., 2022)。具有焦虑倾向的个体，其威胁注意偏向和高警觉性与杏仁核过度活跃有关(Oler et al., 2010; Fox et al., 2018)。而杏仁核的过度激活可能导致个体对线索关联的感知出现紊乱，可能是负性偏向的原因(Stout et al., 2017)。总之，焦虑个体在早期的信息加工处理中，杏仁核和纹状体在面对威胁和奖赏刺激时表现出不平衡的反应，杏仁核的过度活跃加剧了趋避冲突信息中对威胁信息的感知，而对奖赏敏感的纹状体激活程度的较低，导致对威胁成分产生偏向。

3.2 预期价值和动机比较的失衡

信息进入趋避冲突处理阶段时，焦虑个体过高的威胁敏感性导致 BAS 和 FFFS 两个动机系统的过度激活，促使个体采取回避行为以减少不安和恐惧，导致其回避动机增加(Dickson, 2006; Corr & Cooper, 2016)，而 BAS 的活性可能由于担忧和害怕失败而受到抑制，从而减少趋近的动机(Richey et al., 2019; Sequeira et al., 2022)。另一方面，由于对威胁的注意和解释偏差可能会高估潜在威胁，从而增加回避行为的相对价值和权重，以此确保自己的安全，导致焦虑个体对回避反应的期望值偏高。他们更有可能夸大威胁的成本和可能性，形成过度悲观的预期(Mitte, 2008; Grupe & Nitschke, 2013)，并且难以减轻对未来威胁的期望(Zorowitz et al., 2020; Rief et al., 2022)。这导致他们在各种情况下都存在高水平的回避，而不

是特定于某种强化或过度学习的行为(Loijen et al., 2020; Ball et al., 2022)。Pittig 等(2023)借助期望违背来改变焦虑障碍患者的威胁预期，让威胁呈现后不跟随预期的结果（例如，被试预期在发言时被嘲笑的概率很高，实际上得到的是友好的提问）。发现威胁预期被违背的越多，焦虑的治疗效果越好。Charpentier 等(2017)采用赌博任务发现焦虑患者在决策中表现出增强的风险厌恶，而对奖赏得失的重视程度与对照组没有显著差异，与此相似的是，焦虑障碍组可以依靠更少的观察来估计某种行为的厌恶值，而对奖赏价值的估计没有组间差异(Aylward et al., 2019)。综上所述，焦虑个体可能会高估潜在威胁，从而增加回避的预期价值以保证自己的安全，但焦虑对趋近价值的计算所产生的影响还尚未清楚。对焦虑个体的趋避冲突失调的解释需要同时关注动机和价值的异常计算，焦虑个体回避动机的过度激活以及对回避价值的增加，使得其乘积的异常增加，因此即使威胁很小，也表现出过度的回避行为；而趋近动机又因为恐惧和担忧受到抑制，即使对趋近价值的评估与健康个体相同的情况下，其趋近动机和价值的乘积依然低于正常值，因此即使潜在奖赏很大，在进行比较后，焦虑个体依旧选择回避行为。

动机和预期价值的失衡表明，焦虑个体评估趋避冲突以及期望计算的神经回路可能存在异常。当趋避冲突发生时，腹侧海马(vHPC)负责仲裁和比较，在趋近和回避之间进行决策，而 vHPC 对焦虑非常敏感(Bryant & Barker, 2020; Xia & Kheirbek, 2020; Shi et al., 2023)，一方面焦虑可以通过改变突触可塑性，影响 vHPC 的兴奋-抑制平衡，破坏趋避冲突中的比较功能(Couch et al., 2021)；另一方面，vHPC 是调节焦虑的神经回路的中心，其中 vHPC-AMY 回路负责接收与威胁相关的感官刺激，并在处理与焦虑相关的事件中发挥作用，其信号增强导致焦虑行为，焦虑倾向个体 vHPC-AMY 的信息传递同步率显著增强(Felix-Ortiz et al., 2013; Jimenez et al., 2018)。vHPC-mPFC 作为主要的门控回路，可以同时促进或抑制焦虑相关的行为表现，从而实现双向调节，并在 theta 频率范围内传递焦虑信号，这些信息与趋避冲突测试中的回避相关(Padilla-Coreano et al., 2016; Sánchez-Bellot et al., 2022)。总之，vHPC 和相关脑区神经回路的连接是调节焦虑行为的生理基础，在以自上而下的控制占据主导时，由于焦虑导致 vHPC 比较功能的损坏以及在回路中放大焦虑信号，从而让个体对无害的刺激也产生过高的威胁预期。

期望值的计算是一个动态的过程，内部期望主要通过自上而下的神经投射传递，该过程的功能障碍可能导致个体反复对可预测的事件感到意外，夸大了刺激的显著性，导致神经资源过度分配(Howlett & Paulus, 2020)。自上而下的预期编码需要 PFC 的参与，而焦虑个体的威胁期望过高，可能是 PFC 过度的激活从而在权衡时产生的偏差(Mack et al., 2023)。此外，

焦虑还会引发内部预测信号的失调，这主要依赖于纹状体对奖励幅度和概率的处理(Garrison et al., 2013; Cornwell et al., 2017; Bech et al., 2023)，以及杏仁核对威胁信息过度反应所引发的高度预期(Costa et al., 2016; Iordanova et al., 2021)。腹侧纹状体和杏仁核以不同的方式处理信号，然后传入 mPFC 来编码回避和趋近的相对价值和权重。这些系统的稳态失衡将导致期望价值更消极，从而引起趋避冲突失调。

3.3 反馈学习的异常

焦虑个体放大的威胁预期和对回避过高的预期价值会影响得到反馈后的学习过程，然而焦虑对目标导向和习惯化两个学习系统的影响并不一致。在目标导向的学习中，Pittig 等 (2021) 通过趋避冲突来调节焦虑障碍患者的回避反应，结果表明当奖赏和威胁刺激相竞争时，焦虑障碍组无法减少回避，即使在威胁消失后，其回避行为减少仍然有限。而焦虑状态并不会显著损害个体的目标导向的控制(Gillan et al., 2021)；对于习惯化学习的研究表明，焦虑倾向会促进习惯性行为的转变(Flores et al., 2018; Pittig et al., 2020)，在另外一些研究中，相比于健康对照组，焦虑状态组和焦虑障碍组都没有显示出更强的习惯性倾向，即使他们整体表现出较低的趋近率(Gillan et al., 2021; Roberts et al., 2022; Glück et al., 2023)。这种情况可能是由于焦虑个体在目标导向和习惯化两者的切换中存在缺陷，而不是如 Gillan(2015)假设的“脆弱的目标导向系统导致陷入习惯化”。Howlett(2020)要求焦虑障碍被试在两个无关项的干扰下识别目标刺激，而每 30 个试次后，目标刺激出现的频率和获得奖励的条件都会改变，结果发现焦虑对于符合预期的决策更新没有影响，而是与基于环境情境适应的能力受损有关，这可能导致处理预期和实际不符结果的反应过度。焦虑倾向的个体对稳定环境和动荡环境之间调整更新结果预期的能力较弱(Browning et al., 2015; Lamba et al., 2020)，可能是由于更强的先验信念造成的(Paulus, 2020)。因此，焦虑个体的缺陷可能是由于期望的僵化即习惯化与目标导向的信息更新形成抵抗，在需要对两种学习系统进行频繁的转换时，出现不灵活的问题，导致行为调整的失败。

涉及目标导向和习惯化系统的神经机制主要在额叶皮层和纹状体以及周围神经活动的变化(Banca et al., 2015)，这种异常的神经活动主要基于强迫症的研究(见综述 Fineberg et al., 2018)。mPFC 过度激活与无法更新威胁预期显著相关(Apergis-Schoute et al., 2017)，该区域的功能失调以及与纹状体的功能连接减少导致无法灵活地更新恐惧反应，并持续进行习惯性强迫活动。尽管焦虑状态的诱发涉及到纹状体(Robinson et al., 2013; McDermott et al., 2022)，但 Chavanne 和 Robinson (2021)的元分析认为，焦虑状态和焦虑障碍的重叠脑区在于 mPFC 的异常激活，而非重叠区域（如纹状体）的功能缺陷可能是焦虑状态向焦虑障碍转化的中间

机制。焦虑个体在目标导向和习惯化系统之间的转换不灵活，可能是由于前额皮质节点缺乏自上而下的抑制控制，额叶与尾状核的功能连接的减少降低了认知灵活性，从而阻碍了快速的学习转换。因此，焦虑导致反馈学习异常的关键节点或许在于 mPFC 受损，这也需要后续更多的研究来证明。

4. 总结与展望

本文从理论模型和神经机制两个层面对趋避冲突的认知神经机制进行了总结，并在前人研究的基础上提出了趋避冲突的三阶段模型，试图进一步解释焦虑个体趋避冲突失调的整体机制。该模型强调预期值计算和动机比较在趋避冲突处理中的相互作用，并且在得到反馈后更新判断从而帮助个体解决趋避冲突，这依赖于自上而下相关神经回路的编码和比较功能。焦虑会影响模型的各个阶段以及相应神经系统的稳态，导致个体在趋避冲突中的失衡。然而目前的研究还存在一些尚未解决的问题：

4.1 趋避冲突三阶段模型的进一步验证

趋避冲突三阶段模型的提出主要基于三点：（1）动机、条件化以及强化学习等不同领域的理论整合；（2）健康个体在趋避冲突中的行为和神经反应；（3）焦虑个体在趋避冲突中的异常表现以及焦虑涉及趋避冲突相关的神经机制。但将这三点联系起来的实证研究较少，鉴于趋避冲突情景和焦虑不同表型的复杂性，未来的研究或许可以从以下两个方面进行：

一方面，在趋避冲突的复杂情景下，模型中不同的认知机制可能在实验设计中被混淆。常见的潜在奖赏为金钱或积分，潜在威胁则通常是电击或恐怖图片，这样的设置可能导致对回避和趋近的预期价值并不匹配，而个体对于回避动机和趋近动机也存在主观偏好。如果不能严格控制，可能导致预期编码和内部动机之间的混淆，可以在趋避冲突前增加威胁和奖励的匹配程序(Wong & Pittig, 2022)。此外，由于从经验中学习偶发事件的概率与明确提供概率结果对行为决策的影响不同(Hertwig & Erev, 2009; Baczkowski et al., 2023)，表明不同阶段还可能受到刺激结构（如不确定性）的影响。未来还需要进一步的实证工作来确定这些阶段的相对独立性。

另一方面，未来研究需要评估趋避冲突失调与焦虑的关系。焦虑作为一种指向未来的情绪，主要涉及对未来真实/想象的威胁预期(Grupe & Nitschke., 2013;　Fung et al., 2019)，而趋避冲突的解决需要预测和模拟未来事件，对于潜在奖励和威胁的差异模拟分别使得行为偏向趋近与回避(Gilbert & Wilson, 2007; Moughrabi et al., 2022)。后续的研究需要回答威胁感知的增强、动机和预期的失衡以及反馈学习的异常，是否会增加患焦虑障碍的风险？还是说这些

阶段的受损是长期焦虑所导致的后果？此外，以改变威胁预期为主要目的疗法，实现了对焦虑障碍的有效治疗(Craske et al., 2022; Pittig et al., 2023)，后续还可以继续探索焦虑的改善与这三个阶段的神经反应正常化是否相关。

4.2 趋避冲突三阶段模型的参数化

解决趋避冲突的过程可能涉及到刺激本身、个体自身的学习经验以及个体对奖赏或威胁的敏感性差异(Aupperle & Paulus, 2010; Letkiewicz et al., 2023)。其中的任一因素都可能受到焦虑的影响从而导致趋避冲突的失调，然而其中的具体机制难以通过行为反应直接观察到。计算精神病学(computational psychiatry)使用数学模型(例如，基于贝叶斯定理的模型)来解释无法直接观察到的病理行为的心理和神经生理基础(Smith et al., 2020; Vasilchenko & Chumakov, 2023)。通过对学习和决策过程的参数化，研究发现焦虑个体的趋避冲突失调可以用内部处理计算模型中特定参数的变异来解释，比如敏感性(Yamamori et al., 2023)、悲观信念(Zorowitz et al., 2020)、行动的信心(Smith et al., 2021)等。然而，目前将计算模型的方法应用于理解焦虑中的趋避冲突的研究非常有限，并且难以完整描述刺激感知、预期到行为输出到底如何实现。未来可以尝试将趋避冲突的阶段划分为不同的子模型，通过分层和模块化的方法，不仅可以进一步理解焦虑如何影响个体在复杂情景中的表现，还可以通过调整特定的模型参数（如增加或减少威胁，或改变对奖赏的评价等）来模拟不同焦虑程度下的变化。

4.3 从发展的视角考察焦虑个体的趋避冲突失调

先前的研究主要集中在焦虑个体的威胁敏感性和回避特征上(Katz et al., 2020)，其测试的群体以成年人为主，然而焦虑障碍首次出现的高峰时期在青春期，该时期对情绪高度敏感且具有高度的神经可塑性(Towner et al., 2023)。最近的研究发现青少年的焦虑水平和冒险行为成正比，奖赏敏感性在其中起调节作用，焦虑水平高且奖励敏感性高的青少年表现出更多寻求刺激的行为 (Baker et al., 2022; 李晓明 等, 2022)，成年人的焦虑仅仅与回避和威胁敏感性相关，而青少年的焦虑行为表现更为复杂。原因可能是青春期大脑发育的不平衡，其中负责处理奖励和威胁信息的系统（杏仁核和纹状体）在面对刺激时过度激活，而认知控制和比较系统（前额叶和海马体）尚未发育完全(Baker & Galv´an, 2020; Fernández-Teruel, 2021)，导致该时期对正反馈的渴望和对负反馈的害怕，并且难以通过自上而下的调节来平衡冲突。总之，焦虑青少年在面对冲突时表现出过度趋近或回避两种截然相反的行为，在成年人中却只与过度回避有关。未来研究可以通过纵向追踪和多模态的研究来探索焦虑个体趋避冲突失调的发展轨迹。

参考文献

李小新，郭永玉，芈静，胡小勇. (2014). 威胁敏感性的概念和测量：生理和认知两种取向. *心理科学进展*, *22*(10), 1608–1615.

李晓明，刘小丹，戴婷.(2022).特质焦虑与青少年冒险行为的关系：基于双系统模型的机制探讨. *心理与行为研究*(04), 529–535.

Abivardi, A., Khemka, S., & Bach, D. R. (2020). Hippocampal representation of threat features and behavior in a human approach-avoidance conflict anxiety task. *The Journal of Neuroscience*, *40*(35), 6748–6758.

Amodio, D. M., & Harmon-Jones, E. (2011). Trait emotions and affective modulation of the startle eyeblink: On the unique relationship of trait anger. *Emotion, 11*, 47–51.

Apergis-Schoute, A. M., Gillan, C. M., Fineberg, N. A., Fernandez-Egea, E., Sahakian, B. J., & Robbins, T. W. (2017). Neural basis of impaired safety signaling in obsessive compulsive disorder. *Proceedings of the National Academy of Sciences, 114*(12), 3216–3221.

Arnaudova, I., Kindt, M., Fanselow, M., & Beckers, T. (2017). Pathways towards the proliferation of avoidance in anxiety and implications for treatment. *Behaviour Research and Therapy, 96*, 3–13.

Auerbach, R. P., Pagliaccio, D., Hubbard, N. A., Frosch, I., Kremens, R., Cosby, E., & Pizzagalli, D. A. (2022). Reward-Related neural circuitry in depressed and anxious adolescents: A human connectome project. *Journal of the American Academy of Child & Adolescent Psychiatry, 61*(2), 308–320.

Aupperle, R. L., & Paulus, M. P. (2010). Neural systems underlying approach and avoidance in anxiety disorders. *Dialogues in Clinical Neuroscience, 12*(4), 517–531.

Aylward, J., Valton, V., Ahn, W.-Y., Bond, R. L., Dayan, P., Roiser, J. P., & Robinson, O. J. (2019). Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nature Human Behaviour, 3*(10), 1116–1123.

Bach, D. R. (2022). Cross-species anxiety tests in psychiatry: Pitfalls and promises. *Molecular Psychiatry, 27*(1), 154–163.

Bach, D. R., & Dayan, P. (2017). Algorithms for survival: A comparative perspective on emotions. *Nature Reviews. Neuroscience, 18*(5), 311–319.

Bach, D. R., Guitart-Masip, M., Packard, P. A., Miro, J., Falip, M., Fuentemilla, L., & Dolan, R. J. (2014). Human hippocampus arbitrates approach-avoidance conflict. *Current Biology*, *24*(5), 541–547.

Bach, D. R., Hoffmann, M., Finke, C., Hurlemann, R., & Ploner, C. J. (2019). Disentangling hippocampal and amygdala contribution to human anxiety-like behavior. *The Journal of Neuroscience, 39*(43), 8517–8526.

Baczkowski, B. M., Haaker, J., & Schwabe, L. (2023). Inferring danger with minimal aversive experience. *Trends in Cognitive Sciences, 27*(5), 456–467.

Baker, A. E., & Galván, A. (2020). Threat or thrill? The neural mechanisms underlying the development of anxiety and risk taking in adolescence. *Developmental Cognitive Neuroscience, 45*, 100841.

Baker, A. E., Padgaonkar, N. T., Peris, T. S., & Galván, A. (2022). Anxiety symptoms interact with approach motivations in adolescent risk-taking. *https://doi.org/10.31234/osf.io/nqdxc.*

Baladron, J., & Hamker, F. H. (2020). Habit learning in hierarchical cortex-basal ganglia loops. *The European Journal of Neuroscience, 5*2(12), 4613–4638.

Ball, T. M., & Gunaydin, L. A. (2022). Measuring maladaptive avoidance: From animal models to clinical anxiety. *Neuropsychopharmacology*, *47*(5), 978–986.

Banca, P., Voon, V., Vestergaard, M. D., Philipiak, G., Almeida, I., Pocinho, F., Relvas, J., & Castelo-Branco, M. (2015). Imbalance in habitual versus goal directed neural systems during symptom provocation in obsessive-compulsive disorder. *Brain : A Journal of Neurology, 138*(Pt 3), 798–811.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133*(1), 1–24.

Barker, J.M., Bryant, K.G., & Chandler, L.J. (2018). Inactivation of ventral hippocampus projections promotes sensitivity to changes in contingency. *Learning & Memory, 26*, 1 - 8.

Barnett, W. H., Kuznetsov, A., & Lapish, C. C. (2023). Distinct cortico-striatal compartments drive competition between adaptive and automatized behavior. *PloS one, 18*(3), e0279841.

Bech, P., Crochet, S., Dard, R., Ghaderi, P., Liu, Y., Malekzadeh, M., Petersen, C. C. H., Pulin, M., Renard, A., & Sourmpis, C. (2023). Striatal Dopamine Signals and Reward Learning. *Function, 4*(6), zqad056.

Biderman, N., Bakkour, A., & Shohamy, D. (2020). What are memories for? The hippocampus bridges past experience with future decisions. *Trends in Cognitive Sciences, 24*(7), 542–556.

Bishop, S. J., & Gagne, C. (2018). Anxiety, depression, and decision making: A computational perspective. *Annual Review of Neuroscience, 41*, 371–388.

Bolles, R. C. (1970). Species-specific defense reactions and avoidance learning. *Psychological Review, 77*(1), 32–48.

Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience, 18*(4), 590–596.

Bryant, K. G., & Barker, J. M. (2020). Arbitration of Approach-Avoidance Conflict by Ventral Hippocampus. *Frontiers in Neuroscience, 14,* 615337.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333.

Castegnetti, G., Tzovara, A., Khemka, S., Melinščak, F., Barnes, G. R., Dolan, R. J., & Bach, D. R. (2020). Representation of probabilistic outcomes during risky decision-making. *Nature Communications, 11*(1), 2419.

Charpentier, C.J., Aylward, J., Roiser, J.P., & Robinson, O.J. (2017). Enhanced risk aversion, but not loss aversion, in unmedicated pathological anxiety. *Biological Psychiatry, 81*, 1014–1022.

Chavanne, A. V., & Robinson, O. J. (2021). The overlapping neurobiology of induced and pathological anxiety: A meta-analysis of functional neural activation. *The American Journal of Psychiatry, 178*(2), 156–164.

Choi, J. S., & Kim, J. J. (2010). Amygdala regulates risk of predation in rats foraging in a dynamic fear environment. *Proceedings of the National Academy of Sciences of the United States of America, 107*(50), 21773–21777.

Cisler, J. M., Tamman, A. J. F., & Fonzo, G. A. (2023). Diminished prospective mental representations of reward mediate reward learning strategies among youth with internalizing symptoms. *Psychological Medicine, 53*(14), 1–11.

Cornwell, B. R., Garrido, M. I., Overstreet, C., Pine, D. S., & Grillon, C. (2017). The unpredictive brain under threat: a neurocomputational account of anxious hypervigilance. *Biological Psychiatry, 82*(6), 447-454.

Corr, P. J. (2004). Reinforcement sensitivity theory and personality. *Neuroscience & Biobehavioral Reviews, 28*(3), 317–332.

Corr, P. J., & Cooper, A. J. (2016). The reinforcement sensitivity theory of personality questionnaire (RST-PQ): Development and validation. *Psychological Assessment, 28*(11), 1427–1440.

Corr, P. J., & Krupić, D. (2017). Motivating personality: Approach, avoidance, and their conflict. In A. J. Elliot (Ed.), *Advances in motivation science* (pp. 39–90). Elsevier Academic Press.

Corr, P. J., & McNaughton, N. (2012). Neuroscience and approach/avoidance personality traits: A two stage (valuation–motivation) approach. *Neuroscience & Biobehavioral Reviews, 36*(10), 2339–2354.

Costa, V. D., Dal Monte, O., Lucas, D. R., Murray, E. A., & Averbeck, B. B. (2016). Amygdala and ventral striatum make distinct contributions to reinforcement learning. *Neuron, 92*(2), 505-517.

Couch, A. C. M., Berger, T., Hanger, B., Matuleviciute, R., Srivastava, D. P., Thuret, S., & Vernon, A. C. (2021). Maternal immune activation primes deficiencies in adult hippocampal neurogenesis. *Brain, Behavior, and Immunity, 97*, 410–422.

Craske, M. G., Sandman, C. F., & Stein, M. B. (2022). How can neurobiology of fear extinction inform treatment? *Neuroscience & Biobehavioral Reviews, 143*, 104923.

De Pascalis, V., Scacchia, P., Sommer, K., & Checcucci, C. (2019). Psychopathy traits and reinforcement sensitivity theory: Prepulse inhibition and ERP responses. *Biological Psychology, 148*, 107771.

de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal-human translational models. *Psychological Research, 73*(4), 463–476.

Del Arco, A., Park, J., & Moghaddam, B. (2020). Unanticipated stressful and rewarding experiences engage the same prefrontal cortex and ventral tegmental area neuronal populations. *Eneuro, 7*(3), 1-13.

Dickson, J.M., & Dickson, J.M. (2006). Perceived consequences underlying approach goals and avoidance goals in relation to anxiety. *Personality and Individual Differences, 41*(8), 1527–1538.

Diederen, K.M., & Schultz, W. (2015). Scaling prediction errors to reward variability benefits error-driven learning in humans. *Journal of Neurophysiology, 114*(3), 1628–1640.

Diehl, M. M., Bravo-Rivera, C., & Quirk, G. J. (2019). The study of active avoidance: A platform for discussion. *Neuroscience & Biobehavioral Reviews, 107*, 229-237.

Dillon, D. G., Lazarov, A., Dolan, S., Bar-Haim, Y., Pizzagalli, D. A., & Schneier, F. R. (2022). Fast evidence accumulation in social anxiety disorder enhances decision making in a probabilistic reward task. *Emotion, 22*(1), 1–18.

Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D., & Daw, N. D. (2015). Model-based choices involve prospective neural activity. *Nature Neuroscience, 18*(5), 767–772.

Duval, E. R., Javanbakht, A., & Liberzon, I. (2015). Neural circuits in anxiety and stress disorders: A focused review. *Therapeutics and Clinical Risk Management, 11*, 115–126.

Elliot, A. J. (2006). The hierarchical model of approach-avoidance motivation. *Motivation and Emotion, 30*, 111–116.

Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H. J., Ornelas, S., ... & Witten, I. B. (2019). Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature, 570*(7762), 509-513.

Enkhtaivan, E., Nishimura, J., & Cochran, A. (2023). Placing Approach-Avoidance Conflict Within the Framework of Multi-objective Reinforcement Learning. *Bulletin of Mathematical Biology, 85*(11), 116.

Feigley, D. A., & Spear, N. E. (1970). Effect of age and punishment condition on long-term retention by the rat of active- and passive-avoidance learning. *Journal of Comparative and Physiological Psychology, 73*(3), 515–526.

Felix-Ortiz, A. C., Beyeler, A., Seo, C., Leppla, C. A., Wildes, C. P., & Tye, K. M. (2013). BLA to vHPC inputs modulate anxiety-related behaviors. *Neuron, 79*(4), 658–664.

Fernández-Teruel A. (2021). Conflict between Threat Sensitivity and Sensation Seeking in the Adolescent Brain: Role of the Hippocampus, and Neurobehavioural Plasticity Induced by Pleasurable Early Enriched Experience. *Brain Sciences, 11*(2), 268.

Fernández-Teruel, A., & McNaughton, N. (2023). Post-encounter freezing during approach-avoidance conflict: The role of the hippocampus. *Nature Reviews Neuroscience, 24*(7), 451–452.

Fernández-Teruel, A., & Tobeña, A. (2020). Revisiting the role of anxiety in the initial acquisition of two-way active avoidance: pharmacological, behavioural and neuroanatomical convergence. *Neuroscience & Biobehavioral Reviews, 118*, 739–758.

Fineberg, N. A., Apergis-Schoute, A. M., Vaghi, M. M., Banca, P., Gillan, C. M., Voon, V., ... & Robbins, T. W. (2018). Mapping compulsivity in the DSM-5 obsessive compulsive and related disorders: cognitive domains, neural circuitry, and treatment. *International Journal of Neuropsychopharmacology, 21*(1), 42-58.

Flores, A., López, F. J., Vervliet, B., & Cobos, P. L. (2018). Intolerance of uncertainty as a vulnerability factor for excessive and inflexible avoidance behavior. *Behaviour Research and Therapy, 104*, 34–43.

Fox, A. S., Oler, J. A., Birn, R. M., Shackman, A. J., Alexander, A. L., & Kalin, N. H. (2018). Functional connectivity within the primate extended amygdala is heritable and associated with early-life anxious temperament. *Journal of Neuroscience, 38*(35), 7611-7621.

Fung, B. J., Qi, S., Hassabis, D., Daw, N., & Mobbs, D. (2019). Slow escape decisions are swayed by trait anxiety. *Nature Human Behaviour, 3*(7), 702–708.

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 37*(7), 1297–1310.

Gazit, T., Gonen, T., Gurevitch, G., Cohen, N., Strauss, I., Zeevi, Y., Yamin, H., Fahoum, F., Hendler, T., & Fried, I. (2020). The role of mPFC and MTL neurons in human choice under goal-conflict. *Nature Communications, 11*(1), 3192.

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. *Science, 317*(5843), 1351-1354.

Gillan, C. M., Otto, A. R., Phelps, E. A., & Daw, N. D. (2015). Model-based learning protects against forming habits. *Cognitive, Affective, & Behavioral Neuroscience, 15*, 523–536.

Gillan, C. M., Vaghi, M. M., Hezemans, F. H., van Ghesel Grothe, S., Dafflon, J., Brühl, A. B., ... & Robbins, T. W. (2021). Experimentally induced and real-world anxiety have no demonstrable effect on goal-directed behaviour. *Psychological medicine, 51*(9), 1467-1478.

Glück, V. M., Boschet-Lange, J. M., Pittig, R., & Pittig, A. (2023). Persistence of extensively trained avoidance is not elevated in anxiety disorders in an outcome devaluation paradigm. *Behaviour Research and Therapy, 170*, 104417.

Glück, V. M., Zwosta, K., Wolfensteller, U., Ruge, H., & Pittig, A. (2021). Costly habitual avoidance is reduced by concurrent goal-directed approach in a modified devaluation paradigm. *Behaviour Research and Therapy*, *146*, 103964.

Gray, J. A. (1987). Perspectives on anxiety and impulsivity: a commentary. *Journal of Research in Personality, 21*(4), 493–509.

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of septo-hippocampal system*. Oxford University Press.

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: An integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience, 14*(7), 488–501.

Han, S., Gao, J., Hu, J., Ye, Y., Huang, H., Liu, J., Liu, M., Ai, H., Qiu, J., Luo, Y., & Xu, P. (2023). Disruptions of salience network during uncertain anticipation of conflict control in anxiety. *Asian journal of psychiatry, 88*, 103721.

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences, 13*(12), 517–523.

Hofmann, S. G., & Hay, A. C. (2018). Rethinking avoidance: Toward a balanced approach to avoidance in treating anxiety disorders. *Journal of Anxiety Disorders, 55*, 14–21.

Howlett, J. R., & Paulus, M. P. (2020). Where perception meets belief updating: Computational evidence for slower updating of visual expectations in anxious individuals. *Journal of Affective Disorders, 266*, 633–638.

Hundt, N.E., Brown, L.H., Kimbrel, N.A., Walsh, M.A., Nelson-Gray, R.O., & Kwapil, T.R. (2013). Reinforcement sensitivity theory predicts positive and negative affect in daily life. *Personality and Individual Differences, 54*(3), 350-354.

Iordanova, M. D., Yau, J. O.-Y., McDannald, M. A., & Corbit, L. H. (2021). Neural substrates of appetitive and aversive prediction error. *Neuroscience & Biobehavioral Reviews, 123*, 337–351.

Ito, R., & Lee, A. C. H. (2016). The role of the hippocampus in approach-avoidance conflict decision-making: Evidence from rodent and human studies. *Behavioural Brain Research, 313*, 345–357.

Izadpanah, S., Schumacher, M., Bähr, A., Stopsack, M., Grabe, H. J., & Barnow, S. (2016). A 5-year longitudinal study of the adolescent reinforcement sensitivity as a risk factor for anxiety symptoms in adulthood: Investigating the indirect effect of cognitive emotion regulation. *Personality and Individual Differences*, *100*(95), 68–73.

Jimenez, J. C., Su, K., Goldberg, A. R., Luna, V. M., Biane, J. S., Ordek, G., ... & Kheirbek, M. A. (2018). Anxiety cells in a hippocampal-hypothalamic circuit. *Neuron, 97*(3), 670-683.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 313–327.

Kalisch, R., Gerlicher, A. M. V., & Duvarci, S. (2019). A dopaminergic basis for fear extinction. *Trends in Cognitive Sciences, 23*(4), 274–277.

Katz, B. A., Matanky, K., Aviram, G., & Yovel, I. (2020). Reinforcement sensitivity, depression and anxiety: A meta-analysis and meta-analytic structural equation model. *Clinical psychology review, 77*, 101842.

Kaye, S. A., White, M. J., & Lewis, I. (2018). Young females' attention toward road safety images: An ERP study of the revised reinforcement sensitivity theory. *Traffic Injury Prevention, 19*(2), 201–206.

Kelley, N. J., Hortensius, R., Schutter, D. J., & Harmon-Jones, E. (2017). The relationship of approach/avoidance motivation and asymmetric frontal cortical activity: A review of studies manipulating frontal asymmetry. *International Journal of Psychophysiology, 119*, 19-30.

Keramati, M., Dezfouli, A., & Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Computational Biology, 7*(5), e1002055.

Kim, H. F., & Hikosaka, O. (2015). Parallel basal ganglia circuits for voluntary and automatic behaviour to reach rewards. *Brain, 138*(7), 1776–1800.

Kirlic, N., Young, J., & Aupperle, R. L. (2017). Animal to human translational paradigms relevant for approach avoidance conflict decision making. *Behaviour Research and Therapy, 96*, 14–29.

Kirry, A. J., Twining, R. C., & Gilmartin, M. R. (2020). Prelimbic input to basolateral amygdala facilitates the acquisition of trace cued fear memory under weak training conditions. *Neurobiology of Learning and Memory, 172*, 107249.

Korn, C. W., Vunder, J., Miró, J., Fuentemilla, L., Hurlemann, R., & Bach, D. R. (2017). Amygdala lesions reduce anxiety-like behavior in a human benzodiazepine-sensitive approach–avoidance conflict test. *Biological psychiatry, 82*(7), 522-531.

Kreuze, L. J., Jonker, N. C., Hartman, C. A., Nauta, M. H., & de Jong, P. J. (2020). Attentional bias for cues signaling punishment and reward in adolescents: Cross-sectional and prognostic associations with symptoms of Anxiety and behavioral disorders. *Journal of Abnormal Child Psychology, 48*(8), 1007-1021.

Krypotos, A.-M., Effting, M., Kindt, M., & Beckers, T. (2015). Avoidance learning: a review of theoretical models and recent developments. *Frontiers in Behavioral Neuroscience*, *9*, 189.

Lacey, M. F., & Gable, P. A. (2021). Frontal asymmetry in an approach-avoidance conflict paradigm. *Psychophysiology, 58*(5), e13780.

Lamba, A., Frank, M. J., & FeldmanHall, O. (2020). Anxiety impedes adaptive social learning under uncertainty. *Psychological Science, 31*(5), 592–603.

LeBlanc, K. H., London, T. D., Szczot, I., Bocarsly, M. E., Friend, D. M., Nguyen, K. P., ... & Kravitz, A. V. (2020). Striatopallidal neurons control avoidance behavior in exploratory tasks. *Molecular Psychiatry, 25*(2), 491-505.

LeDoux, J. E., & Pine, D. S. (2016). Using neuroscience to help understand fear and anxiety: A two-system framework. *The American Journal of Psychiatry, 173*(11), 1083–1093.

LeDoux, J. E., Moscarello, J., Sears, R., & Campese, V. (2017). The birth, death and resurrection of avoidance: A reconceptualization of a troubled paradigm. *Molecular Psychiatry, 22*(1), 24–36

Lee, D., Seo, H., & Jung, M. W. (2012). Neural basis of reinforcement learning and decision making. *Annual Review of Neuroscience, 35*(1), 287-308.

Letkiewicz, A. M., Kottler, H. C., Shankman, S. A., & Cochran, A. L. (2023). Quantifying aberrant approach-avoidance conflict in psychopathology: A review of computational approaches. *Neuroscience & Biobehavioral Reviews, 147*, 105103.

Lewin, K. (1935). *A dynamic theory of personality.* McGraw-Hill Higher Education.

Liu, L., Wu, J., Geng, H., Liu, C., Luo, Y., Luo, J., & Qin, S. (2022). Long-term stress and trait anxiety affect brain network balance in dynamic cognitive computations. *Cerebral Cortex, 32*(14), 2957-2971.

Livermore, J. J., Klaassen, F. H., Bramson, B., Hulsman, A. M., Meijer, S. W., Held, L., ... & Roelofs, K. (2021). Approach-avoidance decisions under threat: the role of autonomic psychophysiological states. *Frontiers in Neuroscience, 15*, 621517.

Loijen, A., Vrijsen, J. N., Egger, J. I. M., Becker, E. S., & Rinck, M. (2020). Biased approach-avoidance tendencies in psychopathology: A systematic review of their assessment and modification. *Clinical Psychology Review, 77*, 101825.

Lovibond, P. (2006). Fear and Avoidance: An Integrated Expectancy Model. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), *Fear and learning: From basic processes to clinical implications* (pp. 117–132). American Psychological Association.

Ma, L., Chen, W., Yu, D., & Han, Y. (2020). Brain-wide mapping of afferent inputs to accumbens nucleus core subdomains and accumbens nucleus subnuclei. *Frontiers in Systems Neuroscience, 14*, 15.

Mack, N. R., Deng, S., Yang, S. S., Shu, Y., & Gao, W. J. (2023). Prefrontal cortical control of anxiety: Recent advances. *The Neuroscientist, 29*(4), 488-505.

Mathews, A., & MacLeod, C. (2002). Induced processing biases have causal effects on anxiety. *Cognition & Emotion*, *16*(3), 331–354.

McDermott, T. J., Berg, H., Touthang, J., Akeman, E., Cannon, M. J., Santiago, J., ... & Aupperle, R. L. (2022). Striatal reactivity during emotion and reward relates to approach–avoidance conflict behaviour and is altered in adults with anxiety or depression. *Journal of Psychiatry and Neuroscience, 47*(5), E311-E322.

McNally, G. P. (2021). Motivational competition and the paraventricular thalamus. *Neuroscience & Biobehavioral Reviews, 125*, 193–207.

McNaughton, N., DeYoung, C. G., & Corr, P. J. (2016). Approach/avoidance. In J. R. Absher, & J. Cloutier(Eds.), *Neuroimaging personality and character: Traits and mental states in the brain*(pp. 25-49). Elsevier.

Miletić, S., Boag, R. J., Trutti, A. C., Stevenson, N., Forstmann, B. U., & Heathcote, A. (2021). A new model of decision processing in instrumental learning tasks. *Elife, 10*, e63055.

Miller, N. E. (1948). Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *Journal of Experimental Psychology, 38*(1), 89–101.

Miller, S. M., Marcotulli, D., Shen, A., & Zweifel, L. S. (2019). Divergent medial amygdala projections regulate approach–avoidance conflict behavior. *Nature neuroscience, 22*(4), 565-575.

Mitte, K. (2008). Memory bias for threatening information in anxiety and anxiety disorders: A meta-analytic review. *Psychological Bulletin, 134*(6), 886–911.

Monni, A., Olivier, E., Morin, A. J. S., Belardinelli, M. O., Mulvihill, K., & Scalas, L. F. (2020). Approach and avoidance in Gray's, Higgins', and Elliot's perspectives: A theoretical comparison and integration of approach-avoidance in motivated behavior. *Personality and Individual Differences, 166*, 110163.

Moscarello, J. M., & Maren, S. (2018). Flexibility in the face of fear: Hippocampal–prefrontal regulation of fear and avoidance. *Current Opinion in Behavioral Sciences, 19*, 44-49.

Moughrabi, N., Botsford, C., Gruichich, T. S., Azar, A., Heilicher, M., Hiser, J., ... & Cisler, J. M. (2022). Large-scale neural network computations and multivariate representations during approach-avoidance conflict decision-making. *NeuroImage, 264*, 119709.

Mowrer, O. H. (1940). Anxiety-reduction and learning. *Journal of Experimental Psychology, 27*(5), 497–516.

Nguyen, D., Alushaj, E., Erb, S., & Ito, R. (2019). Dissociative effects of dorsomedial striatum D1 and D2 receptor antagonism in the regulation of anxiety and learned approach-avoidance conflict decision-making. *Neuropharmacology, 146*, 222–230.

Oler, J. A., Fox, A. S., Shelton, S. E., Rogers, J., Dyer, T. D., Davidson, R. J., ... & Kalin, N. H. (2010). Amygdalar and hippocampal substrates of anxious temperament differ in their heritability. *Nature, 466*(7308), 864-868.

Otis, J. M., & Mueller, D. (2017). Reversal of cocaine-associated synaptic plasticity in medial prefrontal cortex parallels elimination of memory retrieval. *Neuropsychopharmacology, 42*(10), 2000-2010.

Padilla-Coreano, N., Bolkan, S. S., Pierce, G. M., Blackman, D. R., Hardin, W. D., Garcia-Garcia, A. L., ... & Gordon, J. A. (2016). Direct ventral hippocampal-prefrontal input is required for anxiety-related neural activity and behavior. *Neuron, 89*(4), 857-866.

Pastor, V., & Medina, J. H. (2021). Medial prefrontal cortical control of reward- and aversion-based behavioral output: Bottom-up modulation. *The European Journal of Neuroscience, 53*(9), 3039–3062.

Patterson, D. A. J. (2020). *The role of the ventral hippocampal-accumbens shell circuit in cued approach-avoidance decision making*. University of Toronto.

Paulus, M. P. (2020). Driven by pain, not gain: Computational approaches to aversion-related decision making in psychiatry. *Biological Psychiatry, 87*(4), 359-367.

Paulus, M. P., & Stein, M. B. (2006). An insular view of anxiety. *Biological Psychiatry, 60*(4), 383–387.

Piantadosi, P. T., Yeates, D. C. M., & Floresco, S. B. (2020). Prefrontal cortical and nucleus accumbens contributions to discriminative conditioned suppression of reward-seeking. *Learning & Memory, 27*(10), 429–440.

Pike, A. C., & Robinson, O. J. (2022). Reinforcement learning in patients with mood and anxiety disorders vs control individuals: A systematic review and meta-analysis. *JAMA psychiatry, 79*(4), 313-322.

Pittig, A., & Dehler, J. (2019). Same fear responses, less avoidance: Rewards competing with aversive outcomes do not buffer fear acquisition, but attenuate avoidance to accelerate subsequent fear extinction. *Behaviour Research and Therapy, 112*, 1–11.

Pittig, A., Boschet, J. M., Glück, V. M., & Schneider, K. (2021). Elevated costly avoidance in anxiety disorders: Patients show little downregulation of acquired avoidance in face of competing rewards for approach. *Depression and Anxiety, 38*(3), 361–371.

Pittig, A., Heinig, I., Goerigk, S., Richter, J., Hollandt, M., Lueken, U., ... & Wittchen, H. U. (2023). Change of threat expectancy as mechanism of exposure-based psychotherapy for anxiety disorders: Evidence from 8,484 exposure exercises of 605 patients. *Clinical Psychological Science, 11*(2), 199-217.

Pittig, A., Hengen, K., Bublatzky, F., & Alpers, G. W. (2018). Social and monetary incentives counteract fear-driven avoidance: Evidence from approach-avoidance decisions. *Journal of Behavior Therapy and Experimental Psychiatry, 60*, 69–77.

Pittig, A., Wong, A. H. K., Glück, V. M., & Boschet, J. M. (2020). Avoidance and its bi-directional relationship with conditioned fear: Mechanisms, moderators, and clinical implications. *Behaviour Research and Therapy, 126*, 103550.

Quartz, S. R. (2009). Reason, emotion and decision-making: risk and reward computation with feeling. *Trends in Cognitive Sciences, 13*(5), 209–215.

Rachman, S., & Hodgson, R. (1974). I. Synchrony and desynchrony in fear and avoidance. *Behaviour Research and Therapy, 12*(4), 311–318.

Ranaldi, R. (2014). Dopamine and reward seeking: the role of ventral tegmental area. *Reviews in the Neurosciences, 25*(5), 621-630.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews. Neuroscience*, *9*(7), 545–556.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations on the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.

Richey, J. A., Brewer, J. A., Sullivan-Toole, H., Strege, M. V., Kim-Spoon, J., White, S. W., & Ollendick, T. H. (2019). Sensitivity shift theory: A developmental model of positive affect and motivational deficits in social anxiety disorder. *Clinical Psychology Review, 72*, 101756.

Rief, W., Sperl, M. F., Braun-Koch, K., Khosrowtaj, Z., Kirchner, L., Schäfer, L., ... & Panitz, C. (2022). Using expectation violation models to improve the outcome of psychological treatments. *Clinical Psychology Review, 98*, 102212.

Roberts, C., Apergis-Schoute, A. M., Bruhl, A., Nowak, M., Baldwin, D. S., Sahakian, B. J., & Robbins, T. W. (2022). Threat reversal learning and avoidance habits in generalised anxiety disorder. *Translational Psychiatry, 12*(1), 216.

Robinson, O. J., Overstreet, C., Charney, D. R., Vytal, K., & Grillon, C. (2013). Stress increases aversive prediction error signal in the ventral striatum. *Proceedings of the National Academy of Sciences of the United States of America, 110*(10), 4129–4133.

Roesch, M. R., & Olson, C. R. (2004). Neuronal activity related to reward value and motivation in primate frontal cortex. *Science, 304*(5668), 307–310.

Rozeske, R. R., Jercog, D., Karalis, N., Chaudun, F., Khoder, S., Girard, D., ... & Herry, C. (2018). Prefrontal-periaqueductal gray-projecting neurons mediate context fear discrimination. *Neuron, 97*(4), 898-910.

Rusconi, F., Rossetti, M. G., Forastieri, C., Tritto, V., Bellani, M., & Battaglioli, E. (2022). Preclinical and clinical evidence on the approach-avoidance conflict evaluation as an integrative tool for psychopathology. *Epidemiology and Psychiatric Sciences, 31*, e90.

Sánchez-Bellot, C., AlSubaie, R., Mishchanchuk, K., Wee, R. W., & MacAskill, A. F. (2022). Two opposing hippocampus to prefrontal cortex pathways for the control of approach and avoidance behaviour. *Nature Communications, 13*(1), 339.

Schlund, M. W., Brewer, A. T., Magee, S. K., Richman, D. M., Solomon, S., Ludlum, M., & Dymond, S. (2016). The tipping point: Value differences and parallel dorsal–ventral frontal circuits gating human approach–avoidance behavior. *Neuroimage, 136*, 94-105.

Schumacher, A., Villaruel, F. R., Ussling, A., Riaz, S., Lee, A. C., & Ito, R. (2018). Ventral hippocampal CA1 and CA3 differentially mediate learned approach-avoidance conflict processing. *Current Biology, 28*(8), 1318-1324.

Seligman, M., & Johnston, J. (1973). A cognitive theory of avoidance learning. In F. J. McGuigan, & D. B. Lumsden (Eds.). Contemporary approaches to conditioning and learning (pp. 69–110). Wiley.

Sequeira, S. L., Forbes, E. E., Hanson, J. L., & Silk, J. S. (2022). Positive valence systems in youth anxiety development: A scoping review. *Journal of Anxiety Disorders, 89*, 102588.

Sharp, P. B., Russek, E. M., Huys, Q. J. M., Dolan, R. J., & Eldar, E. (2022). Humans perseverate on punishment avoidance goals in multigoal reinforcement learning. *ELife, 11*, e74402.

Shi, H. J., Wang, S., Wang, X. P., Zhang, R. X., & Zhu, L. J. (2023). Hippocampus: molecular, cellular, and circuit features in anxiety. *Neuroscience bulletin, 39*(6), 1009-1026.

Silvetti, S., & Verguts, T. (2012). Reinforcement learning, high-level cognition, and the human brain. In P. Bright (Ed.), *Neuroimaging:Cognitive and clinical neuroscience* (pp. 283–296). InTech.

Skinner, B. F. (1938). *The behavior of organisms: an experimental analysis.* Appleton-Century.

Smith, R., Kirlic, N., Stewart, J. L., Touthang, J., Kuplicki, R., Khalsa, S. S., ... & Aupperle, R. L. (2021). Greater decision uncertainty characterizes a transdiagnostic patient sample during approach-avoidance conflict: A computational modelling approach. *Journal of Psychiatry and Neuroscience, 46*(1), E74-E87.

Smith, R., Kuplicki, R., Feinstein, J., Forthman, K. L., Stewart, J. L., Paulus, M. P., ... & Khalsa, S. S. (2020). A Bayesian computational model reveals a failure to adapt interoceptive precision estimates across depression, anxiety, eating, and substance use disorders. *PLoS computational biology, 16*(12), e1008484.

Stelly, C. E., Tritley, S. C., Rafati, Y., & Wanat, M. J. (2020). Acute stress enhances associative learning via dopamine signaling in the ventral lateral striatum. *Journal of Neuroscience, 40*(22), 4391-4400.

Stout, D. M., Shackman, A. J., Pedersen, W. S., Miskovich, T. A., & Larson, C. L. (2017). Neural circuitry governing anxious individuals' mis-allocation of working memory to threat. *Scientific Reports, 7*(1), 8742.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine learning, 3*, 9–44.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (Second edition). MIT Press.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*(4), 189–208.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science, 315*(5811), 515–518.

Towner, E., Chierchia, G., & Blakemore, S. J. (2023). Sensitivity and specificity in affective and social learning in adolescence. *Trends in Cognitive Sciences, 27*(7), 642–655.

Tymula, A., Wang, X., Imaizumi, Y., Kawai, T., Kunimatsu, J., Matsumoto, M., & Yamada, H. (2023). Dynamic prospect theory: Two core decision theories coexist in the gambling behavior of monkeys and humans. *Science Advances, 9*(20), eade7972.

Vasilchenko, K. F., & Chumakov, E. M. (2023). Current status, challenges and future prospects in computational psychiatry: A narrative review. *Consortium Psychiatricum, 4*(3), 33–42.

Verharen, J. P., Adan, R. A., & Vanderschuren, L. J. (2020). How reward and aversion shape motivation and decision making: a computational account. *The Neuroscientist, 26*(1), 87-99.

Vogel, S., & Schwabe, L. (2019). Stress, aggression, and the balance of approach and avoidance. *Psychoneuroendocrinology, 103*, 137–146.

Walters, C. J., & Redish, A. D. (2018). A case study in computational psychiatry: Addiction as failure modes of the decision-making system. In A. Anticevic & J. D. Murray (Eds.), *Computational psychiatry: Mathematical modeling of mental illness* (pp. 199–217). Elsevier Academic Press.

Warr, P. B., Sánchez-Cardona, I., Taneva, S. K., Vera, M., Bindl, U. K., & Cifre, E. (2021). Reinforcement sensitivity theory, approach-affect and avoidance-affect. *Cognition and Emotion, 35*(4), 619-635.

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3*(1), 1–14.

Watson, P., O'Callaghan, C., Perkes, I., Bradfield, L., & Turner, K. (2022). Making habits measurable beyond what they are not: A focus on associative dual-process models. *Neuroscience & Biobehavioral Reviews, 142*, 104869.

Wise, T., Liu, Y., Chowdhury, F., & Dolan, R. J. (2021). Model-based aversive learning in humans is supported by preferential task state reactivation. *Science Advances, 7*(31), eabf9616.

Wong, A. H. K., & Pittig, A. (2022). A dimensional measure of safety behavior: A non-dichotomous assessment of costly avoidance in human fear conditioning. *Psychological Research, 86*(1), 312–330.

Wood, W., & Neal, D. T. (2007). A new look at habits and the habit-goal interface. *Psychological Review, 114*(4), 843–863.

Wood, W., & Rünger, D. (2016). Psychology of Habit. *Annual Review of Psychology, 67*(1), 289–314.

Xia, F., & Kheirbek, M. A. (2020). Circuit-based biomarkers for mood and anxiety disorders. *Trends in Neurosciences*, 43(11), 902-915.

Yamamori, Y., & Robinson, O. J. (2023). Computational perspectives on human fear and anxiety. *Neuroscience & Biobehavioral Reviews, 144*, 104959.

Yamamori, Y., Robinson, O. J., & Roiser, J. P. (2023). Approach-avoidance reinforcement learning as a translational and computational model of anxiety-related avoidance. *ELife, 12*, RP87720.

Ye, L., Allen, W. E., Thompson, K. R., Tian, Q., Hsueh, B., Ramakrishnan, C., ... & Deisseroth, K. (2016). Wiring and molecular features of prefrontal ensembles representing distinct experiences. *Cell, 165*(7), 1776-1788.

Yeates, D. C. M., Ussling, A., Lee, A. C. H., & Ito, R. (2020). Double dissociation of learned approach-avoidance conflict processing and spatial pattern separation along the dorsoventral axis of the dentate gyrus. *Hippocampus, 30*(6), 596–609.

Yeates, D. C., Leavitt, D., Sujanthan, S., Khan, N., Alushaj, D., Lee, A. C., & Ito, R. (2022). Parallel ventral hippocampus-lateral septum pathways differentially regulate approach-avoidance conflict. *Nature Communications, 13*(1), 3349.

Zorowitz, S., Momennejad, I., & Daw, N. D. (2020). Anxiety, avoidance, and sequential evaluation. *Computational Psychiatry, 4*,1–17.

Understanding approach-avoidance conflict dysregulation in anxiety: Cognitive processes and neural mechanisms

XIA Yi, ZHANG Jie, ZHANG Huoyin, LEI Yi, DOU Haoran

Institute of Brain and Psychological Sciences, Sichuan Normal University, Chengdu 610066, China

Abstract: Effectively resolving approach-avoidance conflicts is crucial in everyday life. However, anxious individuals exhibit behavioral manifestations of dysregulated approach-avoidance conflict. This dysregulation is characterized by abandoning positive outcomes to avoid stimuli that are unrelated to actual threats or less threatening. Traditional motivational theories divide individuals' coping with approach-avoidance conflict into information input and behavioral output processes. However, these are insufficient to fully explain the specific mechanisms underlying approach-avoidance conflict dysregulation. In this review, we propose a three-stage model comprising conflict perception, conflict processing, and feedback learning. This model emphasizes that approach-avoidance conflict dysregulation in anxious individuals may manifest as heightened threat perception, imbalanced motivation-expected value comparison, and abnormal feedback learning. Future research can further validate the relative independence of these three stages in the model, parameterize the model through hierarchical and modular methods, and explore the mechanisms underlying approach-avoidance conflict dysregulation in anxious individuals through a developmental perspective.

Key words: Anxiety, Approach-Avoidance conflict, expected value, motivation, Cognitive neural mechanism